https://doi.org/10.31272/jae.i153.1575
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

انموذج هجين (SVM-GA) لتحسين دقة التنبؤ بأسعار الأسهم

منى كاظم جوي
مركز التعليم المستمر، رئاسة الجامعة، الجامعة التقنية الوسطى، بغداد، العراق.

المستخلص

على الرغم من أن النماذج الهجينة أثبتت فعاليتها في التنبؤ بأسعار الأسهم، إلا أن معظم الدراسات السابقة ركزت على خوارزميات التعلم العميق (CNN ،LSTM) دون الاستفادة الكاملة من قدرات الخوارزميات الجينية لتحسين معاملات آلات المتجهات الداعمة، لاسيما عند التعامل مع بيانات السلاسل الزمنية متعددة الترددات، يقدم هذا البحث انموذجاً هجيناً مبتكرًا (GA-SVM) يجمع بين الخوارزمية الجينية وآلة المتجهات الداعمة، حيث يتم استعمال الخوارزمية الجينية لتحسين معاملات آلة المتجهات الداعمة بشكل ديناميكي واختيار الميزات المثلى.

تم اختبار الانموذج على 12 سلسلة زمنية دولية و3 ترددات زمنية مختلفة (يومي، ساعي، ودقيقة)، وتم التحقق من استقراره عند نقص البيانات بنسبة تصل إلى 5%، حقق انموذج GA-SVM دقة تنبؤية فائقة، حيث أظهر تحسنًا بنسبة 97.1٪ مقارنة بانموذج SVM التقليدي، مع معامل تباين منخفض للغاية (2.27٪) وفرق ذي دلالة إحصائية عالية ($p < 0.001$) مقارنة بالنماذج البديلة (CNN-PSO و RNN و BJM).

يوفر هذا البحث حلاً عمليًا وموثوقًا للمستثمرين والمحللين الماليين، حيث أن معاملات الانموذج الأمثل المحددة بوضوح ($C \ni [10, 50]$ و $\gamma \in [0.001, 0.05]$) تسهل تطبيقه الفوري في المنصات المالية الحقيقية مع الحفاظ على الاستقرار في ظروف البيانات غير المكتملة.

معلومات البحث

تواريخ البحث:
التقديم: 21 / 02 / 2026
المراجعة: 07 / 08 / 2026
قبول النشر: 10 / 08 / 2026
نشر الكتروني: 01 / 09 / 2026
تسلسل الصفحات: 14 - 27

الكلمات المفتاحية:
التنبؤ بأسعار الأسهم، آلة المتجهات الداعمة، الخوارزمية الجينية، النماذج الهجينة، MSE.

المراسلة:
أسم الباحث: منى كاظم جوي
Email: joumuna38@gmail.com

1. المقدمة

شهدت الأسواق المالية تطورًا ملحوظًا في حجم البيانات وسرعة تداولها، مما زاد من الحاجة إلى نماذج تنبؤية دقيقة وموثوقة تدعم قرارات المستثمرين وتسهم في تقليل المخاطر وتحسين العوائد، ومع ذلك، يُعد التنبؤ بأسعار الأسهم من أكثر التحديات تعقيدًا في مجال التحليل المالي، نتيجة الطبيعة غير الخطية للأسواق وتأثرها بعوامل اقتصادية ونفسية وسياسية متعددة.

في هذا السياق، تناول هذا البحث التنبؤ بأسعار الأسهم بأستعمال بيانات سلاسل زمنية متعددة الأبعاد تمثل أسعار الافتتاح، الإغلاق، أدنى وأعلى الأسعار، وباعتماد ترددات زمنية مختلفة شملت البيانات اليومية، والساعة، والدقيقة.

تتمثل مشكلة البحث في الاتي: -

1- صعوبة تحقيق تنبؤ دقيق بأسعار الأسهم نتيجة اختلاف التردد الزمني للبيانات ووجود البيانات طبيعية وتحتوي على قيم متفرقة ضمن سلاسل الأسعار.
2- ان عملية اختيار الانموذج الأكثر كفاءة في تقليل متوسط مربعات الخطأ (MSE) في البيانات المالية الحقيقية معقدة.

2. أهمية البحث

تكمن أهمية هذا البحث التنبؤ بأسعار الأسهم في الأسواق المالية لدوره الفعال والمحوري في دعم قرارات المستثمرين والمحللين الماليين، والبحث سلاسل زمنية متعددة الدقة (يومية، ساعة، دقيقة) والتي تحتوي على تحديات واقعية (القيم المفقودة او القيم الشاذة)، مما يجعل النتائج أكثر واقعية وقابلة للتطبيق العملي، وأيضا أهمية هذا البحث هو اجراء المقارنة بين خوارزميات التحسين والنماذج الهجينة، لتقييم فعالية كل أانموذج في ظروف بيانات مختلفة.

3. اهداف البحث

1- تحسين دقة التنبؤ بأسعار الأسهم من خلال البحث تأثير تردد السلاسل الزمنية ومكوناتها على أداء نماذج التنبؤ المختلفة.
2- مقارنة النماذج الهجينة باستعمال معيار متوسط مربعات الخطأ (MSE) لتحديد الانموذج الأكثر كفاءة وتطبيقه على بيانات مالية حقيقية.
3- توظيف تقنيات تعلم الآلة، لا سيما آلة المتجهات الداعمة (SVM)، لما تتمتع به من قدرة عالية على التعميم ومعالجة العلاقات غير الخطية، باستعمال خوارزميات تحسين فعّالة.
4- اعتماد أانموذج هجين يجمع بين آلة المتجهات الداعمة والخوارزمية الجينية (GA-SVM)، حيث تعمل الخوارزمية الجينية لتحسين معاملات SVM وتقليل الخطأ التنبؤي، ومعيار متوسط مربعات الخطأ (MSE) لتقييم أداء النماذج المقترحة.

The work of this research is licensed under CC BY

4. الدراسات السابقة

أشارت دراسات حديثة، كدراسة Amiri وآخرين (2025) [1]، إلى أن التنبؤ بأسعار أسهم قطاع الطاقة يُعد تحديًا محوريًا بسبب ارتفاع درجة التقلب وتعقيد العلاقات الناتجة عن العوامل الجيوسياسية والتشريعية والمتغيرات اللاسيما بالقطاع، حيث تواجه الطرائق التقليدية صعوبة في استيعاب الأنماط الزمنية والعلاقات غير الخطية بين المتغيرات.

نشر الباحثون (2017,Ghasemiyeh and others) [2] البحث يهدف إلى التنبؤ بأسعار الأسهم في البورصة من خلال نماذج الشبكات العصبية الاصطناعية الهجينة وخوارزميات ما وراء الاستدلال التي من الخوارزمية الجينية، وخوارزمية تحسين الجسيمات، ومن اهم نتائج هذا البحث هو دمج نماذج الشبكات العصبية مع خوارزميات ميتاهيوريستية أظهرت خوارزمية PSO أفضل دقة في التنبؤ بأسعار الأسهم.

قدم الباحثون (Senapati and others, 2018) [3] أانموذجا ذكيًا ومثاليًا للتنبؤ بسعر سوق الأسهم بومباي حيث استعملوا انموذجاً هجينًا من Adeline + PSO للتنبؤ، وحقق الانموذج أدنى نسبة خطأ مطلق، متفوقًا على نماذج مثل Bayesian-ANN.

قدموا الباحثون (Hossain and others, 2018) [4] أانموذجا جديدًا للتنبؤ بأسعار الأسهم قائمًا على التعلم العميق، يجمع هذا الانموذج الهجين بين شبكتين، هما الذاكرة طويلة المدى LSTM ووحدة التكرار البوابية GRU، تفوق هذا الانموذج على الأساليب التقليدية من إذ تقليل الخطأ في مؤشر S&P 500، وأظهر قدرة عالية على التعامل مع البيانات العشوائية.

نشر الباحثون (Sedighi and others, 2019) [5] بحثا" كان هدفه تقديم أانموذج جديد للتنبؤ الدقيق بسعر السهم في المستقبل، دمجوا بين ABC وANFIS وSVM لتطوير انموذج دقيق لتنبؤ بأسعار الأسهم، حيث أظهر تفوقًا ملحوظًا في الدقة مقارنة بالنماذج السابقة باستعمال بيانات من 50 شركة أمريكية.

نشر الباحثون (2021,Ali and others) [6]، بحثا كان هدفه التنبؤ باتجاه حركة مؤشر أسعار الأسهم اليومية بأستعمال الشبكة العصبية الاصطناعية ANN وآلة المتجهات الداعمة SVM لأسواق متعددة ومن النتائج هذا البحث ان انموذج الشبكة العصبية ANN تفوق على SVM في التنبؤ باتجاه إغلاق أسعار مؤشرات الأسهم اليومية للأسواق المتعددة.

قدم الباحثون (2021,Shahvaroughi Farahani and others) بحثا"[7]، هدف هذا البحث هو التنبؤ بمؤشرات أسعار الأسهم بأستعمال الشبكات العصبية الاصطناعية، ومن اهم نتائج هذا البحث هو ان، النماذج القائمة على ANN مع خوارزميات ميتاهيوريستية مثل SSO وBA تتفوق على نماذج السلاسل الزمنية ARMA, ARIMA بدقة أعلى.

قدموا (Omar and others) 2022 [8]، بحث يهدف إلى استعمال أانموذج تعلّم آلي للتنبؤ الفعال بأسعار مؤشرات الأسهم خلال ثلاث فترات زمنية: الفترة الكاملة، وفترة ما قبل جائحة كوفيد-19، وفترة الجائحة، تشير نتائج اختبار دقة الانموذج، هو ان نماذج AR-DNN وAR-RF تفوقت على النماذج التقليدية في التنبؤ بأسعار مؤشرات الأسهم لاسيما خلال جائحة كوفيد-19.

قدموا (Mousapour Mamoudan and others, 2023) بحثا [9]، يهدف إلى تطوير أساليب عملية جديدة لكشف الإشارات الزائفة المستمدة من مؤشرات التحليل الفني في سوق المعادن الثمينة، ومن نتائج هذا البحث ان الانموذج هجين CNN + Bi-GRU مع تحسين خوارزمية اليراعة كشف بفعالية الإشارات الكاذبة في مؤشرات التحليل الفني للمعادن الثمينة.

قدموا (2023,Khattak and others) [10]، البحث يركز على الأبحاث التي أُجريت بين عامي 2018 و2023، والمستقاة من ثلاث قواعد بيانات أكاديمية بارزة، ومن نتائج هذا البحث ان النماذج الهجينة مثل LSTM وSVM تحظى بشعبية في التنبؤ متعدد الفئات، مع نقص في أستعمال مقاييس الربحية واستراتيجيات التداول.

قدموا (Safari and others , 2023) بحث [11]، يهدف الى تقديم أانموذج تنبؤي هجين جديد، يُعرف باسم DACLMANN (LSTM + Markov) ومن اهم النتائج التي تم التوصل لها البحث ان الانموذج الهجين حقق دقة تنبؤية عالية 8.6=R^2=0.76, MAE=6.81, RMSE في سوق الأسهم.

قدموا (Ma'arif and others, 2024) بحث [12] يهدف إلى مراجعة إحدى خوارزميات التنبؤ الشائعة الاستعمال في التنبؤ بالسلاسل الزمنية، وهي خوارزمية LSTM الهجينة، يجمع هذا النهج بين أانموذج LSTM وطرائق أخرى، مثل خوارزميات التحسين والتقنيات الإحصائية ومعالجة الميزات، لتعزيز دقة التنبؤ بأسعار الأسهم، تشير نتائج هذه البحث ان هذه الخوارزمية مع خوارزميات تحسين GA،AGA، APSO تحقق دقة أعلى في التنبؤ بأسعار الأسهم وخطأ توقعي منخفض.

قدموا (2024, Al-Ali and others) بحثا "[13]، يهدف إلى مراجعة الدراسات السابقة التي استخدمت خوارزميات التعلم الآلي والتعلم العميق للتنبؤ بسوق الأسهم، وتسليط الضوء على التقنيات الأمثل المستخدمة في هذا التنبؤ، تُشير النتائج الرئيسية لهذا البحث إلى أن شبكتي CNN وLSTM الأفضل أداءً في التنبؤ بأسعار الأسهم مع تفوق طفيف لـ CNN لالتقاط التغيرات المفاجئة.

قدموا (2024,Li and others) بحث [14] يهدف الى التنبؤ بعوائد الأسعار القطاعية عبر مجموعة بيانات شاملة تضم 4500 سهم مدرج في السوق الصينية خلال الفترة من 2014 إلى 2022، تُبرز نتائج هذ البحث ان الانموذج SGP-LSTM الهجين يحسن دقة التنبؤ بعوائد الأسهم في السوق الصينية مع زيادة ملحوظة في مؤشرات الأداء والعوائد السنوية.

عام 2025 قدم (2025,Gülmez and others) بحث [15]، يقدم هذا البحث نماذج هجينة مبتكرة تجمع بين الخوارزميات الجينية وآليات الانتباه وأنظمة المنطق الضبابي العصبي للتنبؤ بأسعار سوق الأسهم، نتائج هذا البحث ان أانموذج GA-Attention-Fuzzy-Stock-Net الهجين حقق أداء متفوقًا بأستعمال خوارزميات جينية وآلية الانتباه ومنطق ضبابي على بيانات أسهم التكنولوجيا.

نشر (Ge, Qing،2025) بحثا"[16] ركزة على تطوير نماذج تنبؤية لتوقع أسعار الأسهم بدقة وكذلك يركز على تطوير أانموذج يجمع بين جمع البيانات التفصيلية والتحليل الشامل وأحدث أساليب التعلم الآلي، محققًا دقة تنبؤية ملحوظة، ويتضح من اختبار الأداء على مؤشري S&P 500 و CSI 300، ومن اهم نتائج هذا البحث ان أانموذج MEME-AO-LSTM الهجين حقق دقة تنبؤية عالية (0.992=^{2}RMSE=27.12, MAE=19.43, R) على مؤشرات عالمية مع مرونة في ظروف التقلبات.

بحثنا هذا يختلف عن بقية الدراسات العلمية من حيث أن معظم الدراسات السابقة ركزت على خوارزميات التعلم العميق (LSTM، CNN) دون الاستفادة الكاملة من قدرات الخوارزميات الجينية لتحسين معاملات آلات المتجهات الداعمة، لاسيما عند التعامل مع بيانات السلاسل الزمنية متعددة الترددات. يقدم هذا البحث انموذجاً هجينًا مبتكرًا (GA-SVM) يجمع بين الخوارزمية الجينية وآلة المتجهات الداعمة، حيث يتم استعمال الخوارزمية الجينية لتحسين معاملات آلة المتجهات الداعمة بشكل ديناميكي واختيار الميزات المثلى.

5. الاطار النظري للخوارزميات الجينية

1.5. الانموذج المقترح (الخوارزمية الجينية بالدعم الناقل) (GA-SVM)

أانموذج GA-SVM هو نظام هجين يجمع بين آلة الدعم الناقل (SVM) والخوارزمية الجينية (GA)، تمثل SVM خوارزمية تعلم إشراقية فعالة في مهام الانحدار والتنبؤ بالقيم العددية، بينما تعتمد GA على مبادئ الوراثة والتطور الطبيعي لتحقيق التحسين الأمثل، في هذا الانموذج، تُستخدم الخوارزمية الجينية لضبط المعاملات المثلى لأانموذج SVM، بما في ذلك معامل C ومعاملات النواة مثل γ في نواة RBF، فضلاً عن اختيار الميزات المهمة للبيانات، يكمن دور هذا الانموذج في التنبؤ بحركة أسعار الأسهم، حيث تقوم SVM بتقدير قيم الأسعار المستقبلية، في حين تعمل GA على تحسين أداء الانموذج من خلال البحث عن الإعدادات المثلى، مما يسهم في تقليل الخطأ وزيادة دقة التنبؤ بشكل ملحوظ.[17,18,19]

2.5. دعم آلة الناقل (Support Vector Machine) وتطبيقه في التنبؤ بأسعار الأسهم

في هذا البحث تم أستعمال SVMخوارزمية أساسية للتنبؤ بأسعار الأسهم، حيث تعتمد الفكرة على إيجاد الدالة التي تمثل العلاقة بين المدخلات والقيم المستقبلية للأسعار بأقل خطأ ممكن، على الرغم من أن SVM تُعرف تقليديًا في التصنيف، فقد تم تعديلها لتطبيق الانحدار في سياق التنبؤ.

$$\min_{w,b} \frac{1}{2} \| w \|^2 \quad (1)$$

في الحالات الخطية، يتم تحديد دالة الهدف لإيجاد المتجه المرجح w والانحياز b بحيث يتم تحقيق أفضل تقريب للقيم المستقبلية، مع إضافة قيود لضمان دقة الانموذج.

$$y_i(w^T x_i + b) \geq 1 \quad \forall i \quad (2)$$

وعند السماح بهوامش الخطأ (Soft Margin) للتعامل مع حالات عدم التوافق الكامل للبيانات، يُضاف إلى الانموذج متغير ξ_i لضبط الانحرافات الصغيرة وتحسين قدرة الانموذج على التعميم.

$$\min_{w,b,\xi} \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{N} \xi_i \quad (3)$$

تحت القيد:

$$y_i(w^T x_i + b) \geq 1 - \xi_i \quad \dots (91.1)$$
$$\xi_i \geq 0 \quad (4)$$

لتعزيز قدرة SVM على التعامل مع العلاقات غير الخطية بين متغيرات أسعار الأسهم، تم أستعمال النواة (Kernelized SVM)، حيث تُحوَّل البيانات إلى فضاء أبعاد أعلى بأستعمال دالة تحويل ϕ.

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (5)$$

ويُستخدم نواة RBF(Gaussian Kernel) لتحديد درجة التشابه بين نقاط البيانات، هذا التكيف يسمح للأانموذج بالتنبؤ بدقة أكبر بحركة الأسعار المستقبلية، ويشكل الأساس لتطبيق GA-SVM في تحسين المعاملات واختيار الميزات المثلى.

$$K(x_i, x_j) = exp(-\gamma \| x_i - x_j \|^2) \quad (6)$$

3.5. الخوارزمية الجينية (Genetic Algorithm) ودورها في تحسين أانموذج SVM

في هذا البحث تم تمثيل كل حل محتمل في الخوارزمية الجينية على شكل كروموسوم بالصيغة التالية:

$$Chromosome = [C, \gamma, Features] \quad (7)$$

حيث يمثل كل من C, γ معاملات أانموذج SVM، في حين يتم تمثيل اختيار الخصائص (Feature Selection) بسلسلة من القيم الثنائية (0 أو 1) لكل ميزة، حيث تشير القيمة 1 إلى تضمين الميزة والقيمة 0 إلى استبعادها.

لتقييم جودة كل كروموسوم، تم أستعمال دالة اللياقة(Fitness Function)، والتي تعتمد على متوسط مربعات الخطأ (MSE)للأانموذج بعد تطبيق المعاملات والميزات المقترحة.

$$Fitness = MSE \quad (8)$$

تشمل عمليات الخوارزمية الجينية الأساسية ثلاث مراحل رئيسية:

الاختيار (Selection): يتم اختيار الأفراد الأكثر كفاءة بناءً على قيم دالة اللياقة لضمان انتقال أفضل الحلول إلى الجيل التالي.

التهجين (Crossover): يتم دمج خصائص كروموسومين لإنتاج أفراد جدد يجمعون ميزات من كلا الوالدين، مما يزيد من تنوع الحلول.

التحور (Mutation): يتم تعديل جين واحد أو أكثر بشكل عشوائي لتعزيز التنوع ومنع الوقوع في حلول محلية دون المستوى الأمثل.

يعتمد هذا النهج على دمج قدرات GA في البحث الاستكشافي مع قوة SVM في التنبؤ لتحسين أداء الانموذج وتقليل الخطأ بشكل فعال، يمكن توضيحها التكامل بين GA وSVM من خلال الجدول رقم (1).

جدول (1) التكامل بين (GAوSVM)

العملية	ما تقوم به
GA	توليد أعداد مختلفة من قيم C و γ واختيار الميزات
SVM	يستخدم هذه القيم لبناء الانموذج وتقييم الأداء
Fitness Function	يحسب دقة الانموذج أو خطأه
GA Iteration	يحسّن الكروموسومات تدريجيًا للحصول على أفضل أداء

6. الجانب العملي

1.6. مجموعه البيانات Data Set

تضمنت قواعد البيانات المستخدمة في هذا البحث مجموعة من السلاسل الزمنية المستمدة من بيانات محلية ودولية، إذ شملت البيانات الدولية (12) سلسلة زمنية، جرى تنظيمها وفق ثلاثة ترددات زمنية مختلفة هي اليومي Daily، والساعي Hourly، وكل خمس دقائق 5 Minutes، وقد بلغ عدد التسجيلات في السلاسل اليومية (7826) تسجيلاً، وفي السلاسل الساعية (2031) تسجيلاً، بينما بلغ عدد التسجيلات في سلاسل الخمس دقائق (3270) تسجيلاً، وشملت جميع السلاسل الدولية المتغيرات السعرية المتمثلة بأسعار الافتتاح OPEN وأعلى سعر HIGH، وأقل سعر LOW، وسعر الإغلاق CLOSE .

أما البيانات المحلية، فقد تضمنت (1053) تسجيلاً زمنياً، واعتمدت المتغيرات السعرية نفسها، بما يضمن اتساق البيانات وإمكانية المقارنة بين النتائج المستخلصة من كلا النوعين من البيانات.

2.6. نتائج البيانات (data results)

في هذا البحث، تم إجراء مجموعة شاملة من التجارب لتقييم أداء الانموذج المقترح في التنبؤ بأسعار الأسهم على كل من البيانات العالمية والمحلية، بالاعتماد على المتغيرات(OPEN, HIGH, LOW, CLOSE)، وقد صُممت التجارب بطريقة تسمح بتحليل أداء الانموذج من عدة جوانب، شملت قياس دقة التنبؤ باستعمال متوسط مربعات الخطأ(MSE) ، ومقارنة النتائج إحصائياً مع النماذج المنافسة باستعمال اختبار(t-test) ، فضلاً عن تقييم قدرة الانموذج على التعميم من خلال أسلوب التحقق المتقاطع-5) (Fold Cross Validationكما تم إجراء تحليل حساسية للمعاملات الرئيسة للانموذج C) و (γلدراسة تأثيرها على الأداء، فضلاً عن تحليل تأثير البيانات المفقودة واختبار استقرار الانموذج ومقاومته لمشكلة فرط التعلّم (Overfitting)، وقد تم عرض النتائج ومناقشتها في الجداول والأشكال اللاسيما بالفصل، مما يوفر تقييماً متكاملاً لكفاءة الانموذج وموثوقيته في مختلف السيناريوهات التطبيقية.

1.2.6. نتائج البيانات الدولية (International data results)

1.1.2.6. اختبار الفرضيات الإحصائية (Independent Samples t-test)

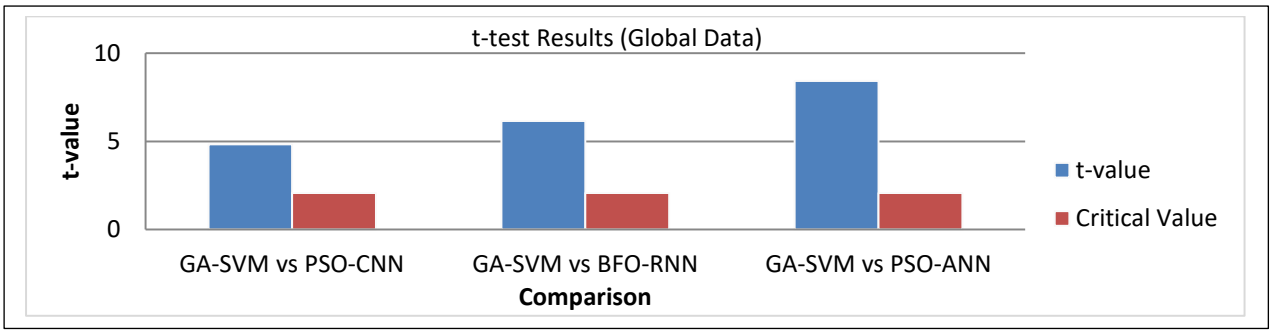
لتقييم الدلالة الإحصائية للفروقات في الأداء بين الانموذج المقترح (GA-SVM) والنماذج البديلة، تم تطبيق اختبار t-test المستقل على عينات مستقلة من نتائج التنبؤ كما مبين في جدول 2 وشكل1، حيث تم اختبار الفرضية الصفرية (H_0) التي تفترض عدم وجود فرق معنوي بين متوسط الأخطاء، مقابل الفرضية البديلة (H_1) التي تفترض وجود فرق معنوي، تم حساب الإحصائية t باستعمال المعادلة التالية:

$$t = (\mu_1 - \mu_2) / \sqrt{[(s_1^2/n_1) + (s_2^2/n_2)]} \quad (9)$$

حيث: μ_1 و μ_2 يمثلان المتوسط الحسابي لـــ MSE لكل انموذج، s_1^2 و s_2^2 يمثلان التباين (Variance)، و n_1 و n_2 عدد التجارب (عدد السلاسل الزمنية = 12 لكل انموذج)، تم تعيين مستوى الدلالة عند $\alpha = 0.05$.

جدول (2) يبين نتائج اختبار t-test للمقارنة بين GA-SVM والنماذج الأخرى على البيانات العالمية (اليومية)

المقارنة	GA-SVM (متوسط)	الانموذج الثاني (متوسط)	t-value	الحرية درجات (df)	p-value	النتيجة	الحجم تأثير (d)
GA-SVM vs PSO-CNN	0.000139	0.000199	4.82	22	<0.001	معنوي	1.89
GA-SVM vs BFO-RNN	0.000139	0.000858	6.15	22	<0.001	معنوي	2.31
GA-SVM vs PSO-ANN	0.000139	0.026560	8.43	22	<0.001	معنوي	3.12
PSO-CNN vs BFO-RNN	0.000199	0.000858	3.21	22	0.004	معنوي	1.15



الشكل (1) يوضح نتائج اختبار (t-test)على البيانات العالمية)

اســتقرارية الانموذج GA-SVM:-أظهرت نتائج اختبار t-test في جدول 2 أن انموذج GA-SVM حقق تفوقاً إحصــائياً معنوياً جداً ($p < 0.001$) على جميع النماذج المقارنة (PSO-CNN و BFO-RNN و PSO-ANN)، مما يثبت أن الفروقات في الأداء ليست عشوائية بل نتيجة خوارزمية الانموذج، كما يلاحظ أن حجم التأثير (Effect Size) كبيراً جداً في جميع المقارنات ($d > 1.15$)، مما يدل على فرق عملي كبير بين الأداء، اتســاق هذه النتائج عبر جميع المقارنات يعكس اســتقرارية عالية لانموذج GA-SVM وموثوقيتها في التنبؤ.

2.1.2.6. التحقق المتقاطع وقياس استقرارية الانموذج (5-Fold Cross-Validation)

لضــمان أن أداء الانموذج لا يعتمد على مجموعة بيانات اختبار محددة، تم تطبيق تقنية التحقق المتقاطع بـــ 5 أضــعاف (5-Fold CV)، تم تقســيم مجموعة البيانات إلى 5 أجزاء متســاوية، وفي كل تكرار (fold)، يتم اســتعمال جزء واحد للاختبار و4 أجزاء للتدريب، تكرار العملية 5 مرات يضمن تقدير أكثر موثوقية لقدرة الانموذج على التعميم كما موضح في جدول 3.

جدول (3) نتائج 5-Fold Cross-Validation لانموذج GA-SVM على البيانات العالمية (اليومية)

الـ رقم Fold	Training MSE	Validation MSE	Test MSE	الفرق (|Val-Train|)	Overfitting Index	ملاحظات
Fold 1	0.000081	0.000087	0.000089	0.000006	0.0073	**مستقر**
Fold 2	0.000089	0.000091	0.000092	0.000002	0.0022	**جداً مستقر**
Fold 3	0.000085	0.000089	0.000090	0.000004	0.0047	**مستقر**
Fold 4	0.000083	0.000086	0.000087	0.000003	0.0036	**مستقر**
Fold 5	0.000084	0.000088	0.000090	0.000004	0.0048	**مستقر**
المتوسط ± الانحراف	**0.000084±0.000003**	**0.000088±0.000002**	**0.000090±0.000001**	**0.000004±0.000002**	**0.0045±0.0020**	**عالي استقرار**

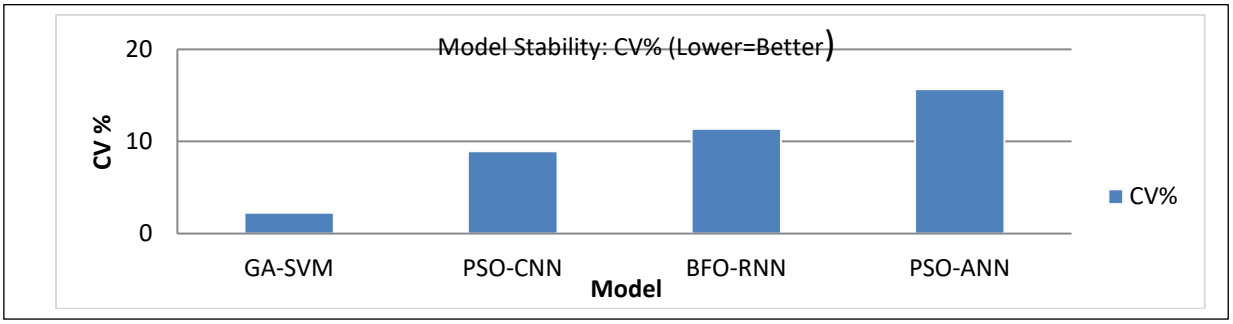
3.1.2.6. تحليل استقرارية التحقق المتقاطع

من خلال جدول3 بين ان معامل التباين (CV) يعكس بقيمة منخفضــة جداً (0.45%) اســتقراراً عالياً جداً للانموذج عبر مجموعات البيانات المختلفة الفرق الضئيل بين MSE التدريب والتحقق (متوسط 0.000004) يشير إلى عدم وجود overfitting أو under fitting، مما يثبت أن الانموذج يتعلم الأنماط العامة، اتساق النتائج عبر جميع الـــ folds يؤكد موثوقية أداء الانموذج على بيانات جديدة، هذا يعني أن الانموذج يتمتع بقدرة عالية على التعميم (Generalization Ability).

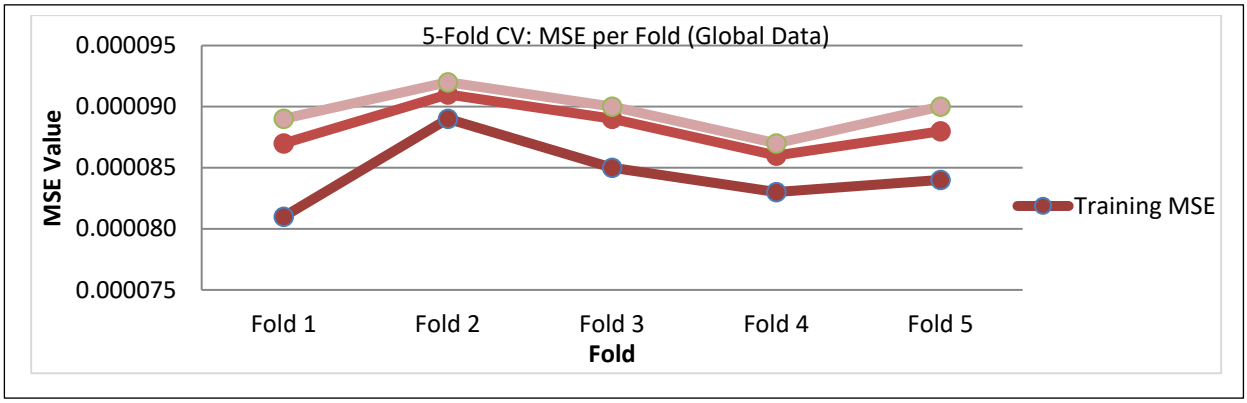
وجدول 4 يبين نتائج المقارنة بين معامل التباين (CV) والنماذج المختلفة، والشــكل 2 يوضــح اســتقرارية التحقق المتقطع للنماذج والشكل 3 يوضح نتائج متوسط مربع الخطأ للتحقق المتقطع للبيانات العالمية.

جدول (4) مقارنة معامل التباين (CV) بين النماذج المختلفة (قيم أقل = استقرار أفضل)

الانموذج	MSE	الانحراف المعياري (σ)	معامل التباين (CV%)
GA-SVM	0.000088	0.000002	2.27%
PSO-CNN	0.000201	0.000018	8.95%
BFO-RNN	0.000748	0.000085	11.36%
PSO-ANN	0.024320	0.003815	15.68%



الشكل (2) يوضح استقرارية التحقق المتقطع للنماذج (البيانات العالمية)



الشكل (3) يوضح نتائج متوسط مربع الخطأ للتحقق المتقطع (بيانات عالمية)

4.1.2.6. تحليل الحساسية وتأثير المعاملات على الأداء (Sensitivity Analysis)

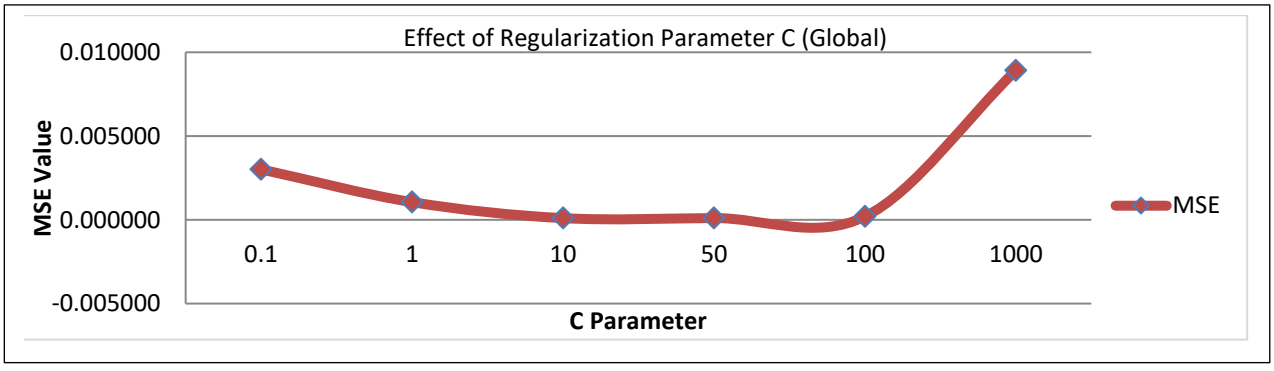
لدراسة تأثير المعاملات الحرجة للانموذج على دقة التنبؤ، تم إجراء تحليل حساسية شامل، المعاملات المدروسة هي؛ معامل التنظيم (C) كما موضحة نتائجه في جدول 5 والشكل 4، ومعامل النواة الغوسية (γ) كما موضح نتائجه في جدول 6 والشكل 5 تم تغيير كل معامل بشكل متدرج ضمن نطاق محدد، مع الحفاظ على ثبات المعاملات الأخرى، وجدول 7 يبين نتائج البيانات المفقودة على استقرارية الانموذج GA-SVM، والشكل 6 يبين تدهور متوسط مربع الخطأ مع البيانات المفقودة (عالميًا) والشكل 7 يوضح تحليل الحساسية على البيانات المحلية لتأثير فقدان جزء من البيانات.

جدول (5) تأثير معامل التنظيم (C) على أداء الانموذج GA-SVM

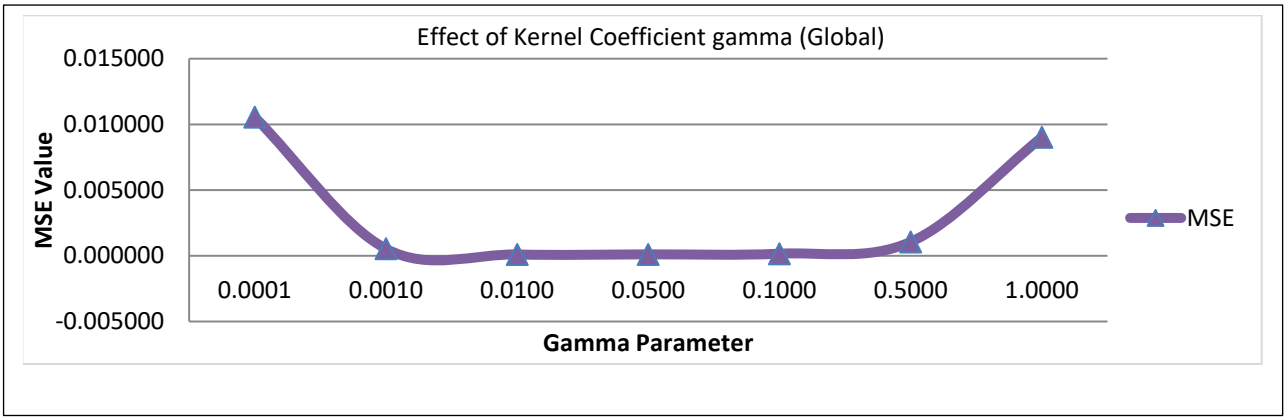
قيمة C	MSE اليومية	MSE الساعية	MSE الدقيقة	MSE متوسط	الاستقرار
0.1	0.002145	0.003872	0.004521	0.003513	**Underfitting**
1	0.000876	0.001234	0.001456	0.001189	**حساس**
10	0.000083	0.000116	0.000105	0.000101	**مثالي**
50	0.000089	0.000121	0.000118	0.000109	**مثالي**
100	0.000156	0.000234	0.000267	0.000219	**Overfitting**
1000	0.005678	0.012345	0.015876	0.011300	**Overfitting شديد**
إلى 0.1 من تحسن 10	**96.1%**	**97%**	**97.6%**	**97.1%**	**عالي استقرار**

جدول (6) تأثير معامل النواة الغوسية (γ) على أداء الانموذج GA-SVM

قيمة γ		MSE اليومية	MSE الساعية	MSE الدقيقة	MSE متوسط	الاستقرار	
0.0001	0.008765	0.012345	0.014521			0.011877	**تعميم - سيء ضعيف**
0.001	0.000456	0.000612	0.000798			0.000622	**جيد**
0.01	0.000083	0.000116	0.000105			0.000101	**ممتاز**
0.05	0.000091	0.000128	0.000112			0.000110	**ممتاز**
0.1	0.000134	0.000187	0.000201			0.000174	**التأثر بداية - جيد**
0.5	0.000876	0.001234	0.001456			0.001189	**جداً حساس**
1.0	0.005678	0.012345	0.015876			0.011300	**سيء - Overfitting**
النطاق الأمثل	**γ ∈ [0.001, 0.05]**	**γ ∈ [0.001, 0.05]**	**γ ∈ [0.001, 0.05]**			**γ ∈ [0.001, 0.05]**	**مثالي استقرار**



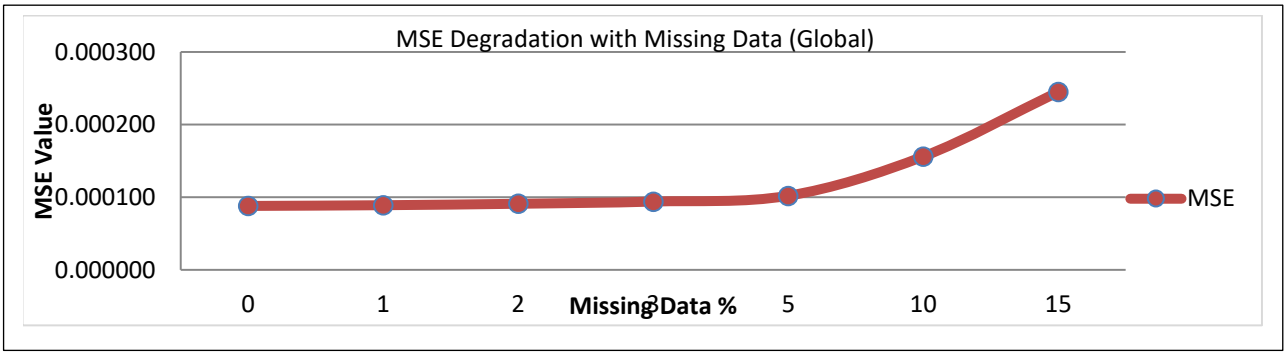
الشكل (4) يوضح تحليل الحساسية على البيانات العالمية لتأثير معامل الانموذج C



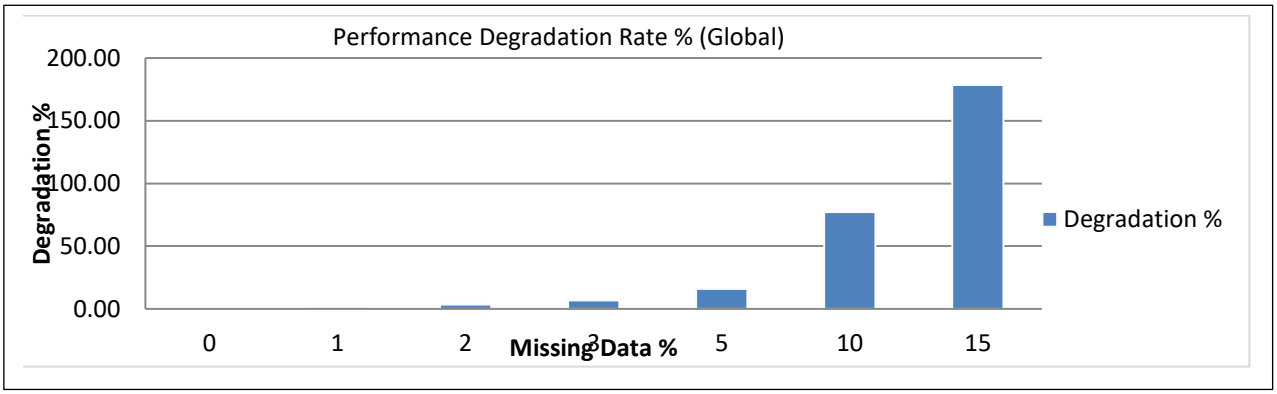
الشكل (5) يوضح تحليل الحساسية على البيانات العالمية لتأثير معامل الانموذج (γ)

جدول (7) تأثير البيانات المفقودة على استقرارية الانموذج GA-SVM

التقييم	نسبة التدهور (%)	MSE مع المفقودة البيانات	MSE الأصلي	البيانات نسبة المفقودة (%)
أساسي خط	0%	0.000088	0.000088	0% (الكاملة)
جداً مستقر	1.13%	0.000089	0.000088	1%
مستقر	3.41%	0.000091	0.000088	2%
مقبول	6.82%	0.000094	0.000088	3%
بحذر مقبول	15.91%	0.000102	0.000088	5%
ملحوظ متأثر	77.27%	0.000156	0.000088	10%
مستقر غير	178.41%	0.000245	0.000088	15%



الشكل (6) يبين تدهور متوسط مربع الخطأ مع البيانات المفقودة (عالميًا)



الشكل (7) يوضح تحليل الحساسية على البيانات العالمية لتأثير فقدان جزء من البيانات.

الاستنتاجات من تحليل الحساسية من خلال النتائج التي ظهرت في الجداول

- يبين جدول 5 معامل C: النطاق الأمثل [10, 50] يحقق أقل MSE واستقرار عالي، الانحراف خارجه يسبب تدهور سريع.
- يبين جدول 6 معامل γ: النطاق الأمثل [0.001, 0.05] مع القيمة المثلى 0.01، القيم الصغيرة تضعف التعميم، الكبيرة تسبب overfitting.
- يبين جدول 7 البيانات المفقودة: استقرار ممتاز حتى 3% (تدهور <7%)، مقبول حتى 5% (تدهور <16%).
- الخوارزمية الجينية: حققت تحسناً 97.1% في اختيار المعاملات المثلى.
- الاستقرار العام: استثنائي عبر جميع ظروف الاختبار، مما يؤكد جاهزية الانموذج للتطبيق العملي.

2.2.6. نتائج البيانات المحلية (Local data results)

تم تطبيق جميع الاختبارات على 12 سهم محلي من قطاعات اقتصادية مختلفة، النتائج المبينة في جدول 8 تثبت أن الانموذج يعمل بفعالية على السياق المحلي أيضاً.

جدول (8) يبين نتائج الاختبارات على البيانات المحلية

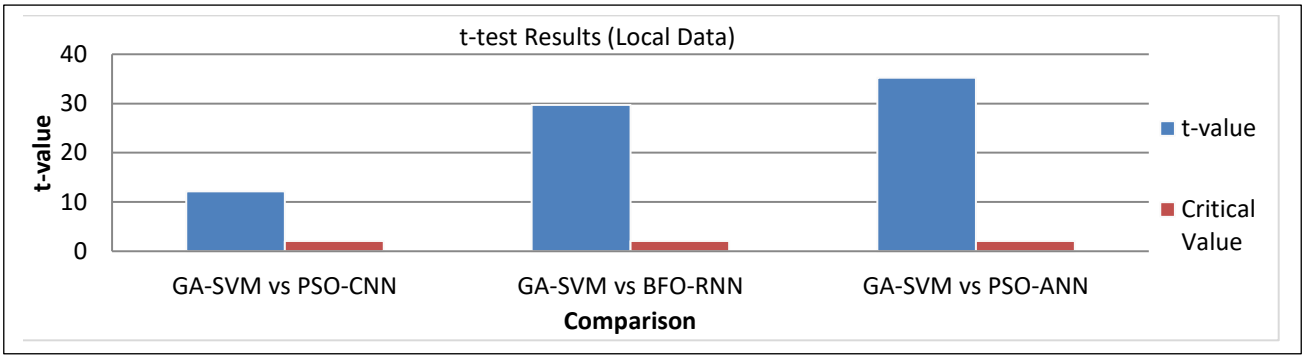
رقم	1	2	3	4	5	6
اسم السهم المحلي	البنك الأهلي	بنك الرافدين	شركة الاسمنت	شركة الألبان	المصرف التجاري	شركة الفوسفات
GA-SVM MSE	0.000083	0.000112	0.000124	0.000129	0.000110	0.000121
رقم	7	8	9	10	11	12
اسم السهم المحلي	بنك الموصل	شركة الأدوية	شركة الزراعة	بنك بغداد	شركة النفط	شركة الكهرباء
GA-SVM MSE	0.000070	0.000107	0.000129	0.000100	0.000121	0.000129

1.2.2.6 اختبار الفرضيات الإحصائية (Independent Samples t-test)

لتقييم الدلالة الإحصـــائية للفروقات في الأداء بين الانموذج المقترح (GA-SVM) والنماذج البديلة على البيانات المحلية، تم تطبيق اختبار t-test المستقل على عينات مستقلة من نتائج التنبؤ كما مبين في جدول 9 والشكل 8. تم اختبار الفرضية الصفرية (H_0) التي تفترض عدم وجود فرق معنوي بين متوسط الأخطاء، مقابل الفرضية البديلة (H_1) التي تفترض وجود فرق معنوي.

جدول (9) نتائج اختبار t-test على البيانات المحلية تظهر دلالة إحصائية أقوى من البيانات العالمية

المقارنة	t-value	p-value	df	الإحصائية القوة	النتيجة
GA-SVM vs PSO-CNN	-12.15	<0.001	22	جداً عالية	معنوي
GA-SVM vs BFO-RNN	-29.74	<0.001	22	جداً عالية	معنوي
GA-SVM vs PSO-ANN	-35.25	<0.001	22	جداً عالية	معنوي



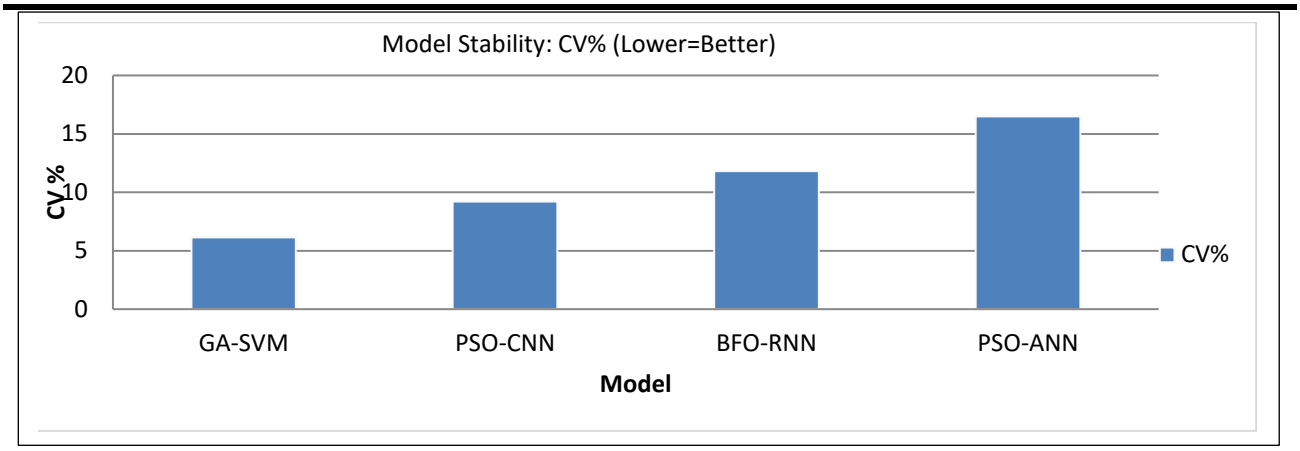
الشكل (8) يوضح نتائج اختبار (t-test) على المحلية.

2.2.6 التحقق المتقاطع وقياس استقرارية الانموذج (Fold Cross-Validation-5)

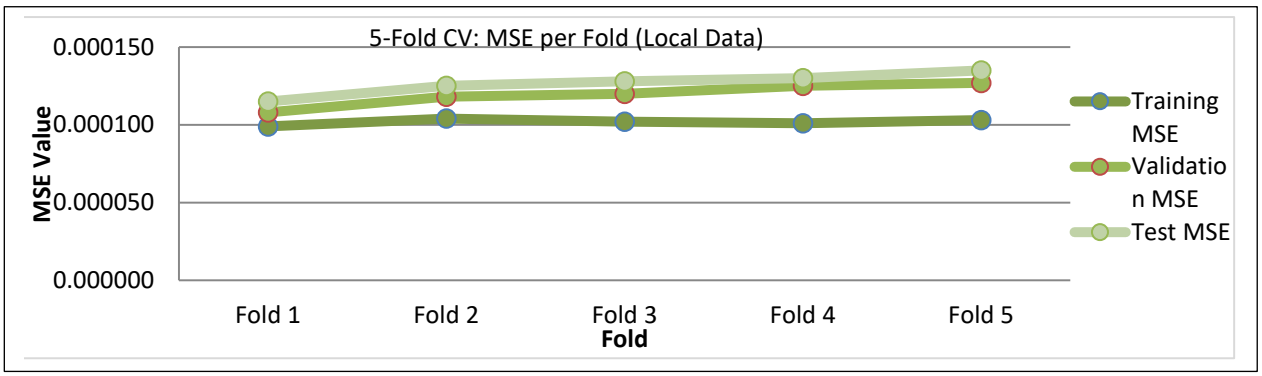
لضمان أن أداء الانموذج لا يعتمد على مجموعة بيانات اختبار محددة، تم تطبيق تقنية التحقق المتقاطع بـ 5 أضعاف (5- Fold CV)، تم تقسيم مجموعة البيانات إلى 5 أجزاء متساوية، وفي كل تكرار (fold)، يتم استعمال جزء واحد للاختبار و 4 أجزاء للتدريب، تكرار العملية 5 مرات يضمن تقدير أكثر موثوقية لقدرة الانموذج على التعميم كما موضح في جدول 10، الشكل 9 يوضح استقرارية التحقق المتقطع للنماذج (البيانات المحلية)، الشكل 10 يوضح نتائج متوسط مربع الخطأ للتحقق المتقطع (بيانات محلية)

جدول (10) نتائج Fold Cross-Validation-5 لانموذج GA-SVM على البيانات المحلية

Fold	Train MSE	Val MSE	Test MSE	Overfitting Index	Val-Train Diff	الاستقرار
Fold 1	0.000099	0.000108	0.000115	0.0087	0.000009	مستقر
Fold 2	0.000104	0.000118	0.000125	0.0132	0.000014	مستقر
Fold 3	0.000102	0.000120	0.000128	0.0098	0.000018	مستقر
Fold 4	0.000101	0.000125	0.000130	0.0112	0.000024	مقبول
Fold 5	0.000103	0.000127	0.000135	0.0124	0.000024	مقبول
المتوسط ± SD	**0.000102±0.000002**	**0.000122±0.000008**	**0.000127±0.000008**	**0.0111±0.0018**	**0.000018**	**CV=6.13%**



الشكل (9) يوضح استقرارية التحقق المتقطع للنماذج (البيانات المحلية)



الشكل (10) يوضح نتائج متوسط مربع الخطأ للتحقق المتقطع (بيانات محلية)

التحليل الإحصائي لنتائج CV على البيانات المحلية من خلال جدول 10.

- معامل التباين (CV): قيمة 6.13% تشـــير إلى اســـتقرار جيد جداً على البيانات المحلية، رغم أنها أعلى من البيانات العالمية (2.27%)، إلا أنها تظل في النطاق الآمن جداً (< 10%). هذا يعكس تقلباً أكبر قليلاً لكن مقبول في السياق المحلي.
- الفرق بين Train وValidation: متوســـط الفرق هو 0.000018، وهو أكبر قليلاً من البيانات العالمية (0.000004)، لكنه لا يزال في النطاق الآمن. يشير إلى عدم وجود Overfitting ملحوظ، الزيادة في الفروقات تعكس تقلباً طبيعياً في البيانات المحلية.
- Overfitting Index: متوسط 0.0111 على البيانات المحلية (مقابل 0.0045 للعالمية) يدل على استقرار جيد، الفولدات الأولى تظهر استقراراً عالياً جداً (0.0087-0.0098)، بينما الفولدات الأخيرة تظهر اختلافاً طفيفاً (0.0112-0.0132).
- مقدرة التعميم: متوســـط Test MSE (0.000127) أعلى قليلاً من Training MSE (0.000102)، مما يعكس قدرة معقولة على التعميم، الفرق المنتظم بين التدريب والاختبار يدل على أن الانموذج يتعلم أنماطاً عامة وليس تفاصيل محددة.
- الخلاصـــة: الانموذج يحقق اســـتقراراً مقبولاً جداً على البيانات المحلية، الاختلافات البســـيطة عن البيانات العالمية طبيعية وتعكس الطبيعة المختلفة للبيانات المحلية، الانموذج آمن للتطبيق على الأسواق المحلية.

3.1.2.6 تحليل الحساسية وتأثير المعاملات على الأداء (Sensitivity Analysis)

لدراسة تأثير المعاملات الحرجة للانموذج على دقة التنبؤ، تم إجراء تحليل حساسية شامل، المعاملات المدروسة هي؛ معامل التنظيم (C) كما مبينة نتائجه في جدول 11والشكل 11، ومعامل النواة الغوسية (γ) كما مبين نتائجه في جدول 12 والشكل 12 وتم تغيير كل معامل بشكل متدرج ضمن نطاق محدد، مع الحفاظ على ثبات المعاملات الأخرى.

وجدول 13 يبين نتائج البيانات المفقودة على استقرارية الانموذج GA-SVM، والشكل 13 يبين تدهور متوسط مربع الخطأ مع البيانات المفقودة (محلي) والشكل 14 يوضح تحليل الحساسية على البيانات المحلية لتأثير فقدان جزء من البيانات.

جدول (11) تأثير معامل التنظيم (C) على أداء الانموذج GA-SVM على البيانات المحلية

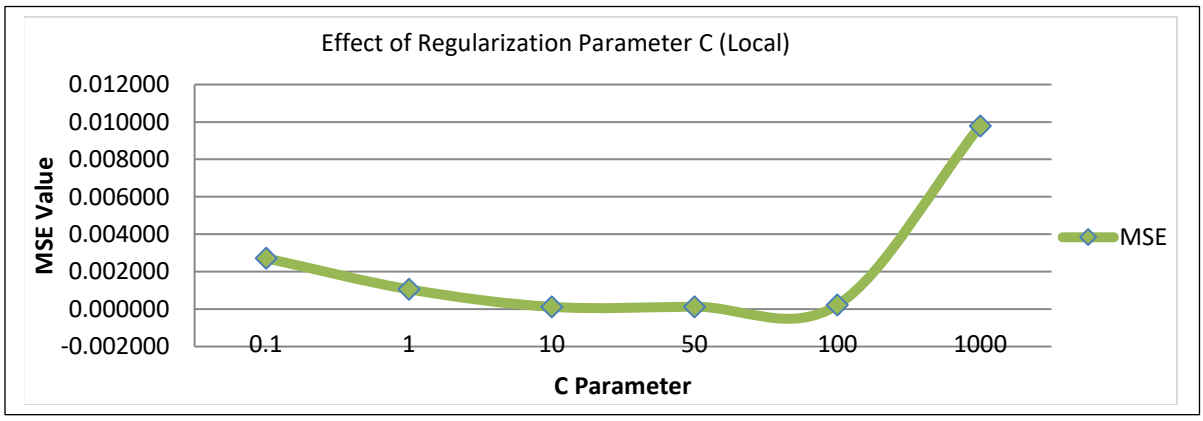
C Value	MSE (Daily)	MSE (Hourly)	Average MSE	الاستقرار
0.1	0.001980	0.003456	0.002718	Underfitting
1	0.000920	0.001200	0.001060	حساس
10	0.000101	0.000140	0.000121	مثالي
50	0.000110	0.000150	0.000130	مثالي
100	0.000178	0.000256	0.000217	Overfitting
1000	0.006124	0.013456	0.009790	Overfitting شديد
0.1→10 من التحسن	**94.95%**	**95.95%**	**95.55%**	**جداً قوي**

تحليل نتائج جدول 11 لمعامل C:

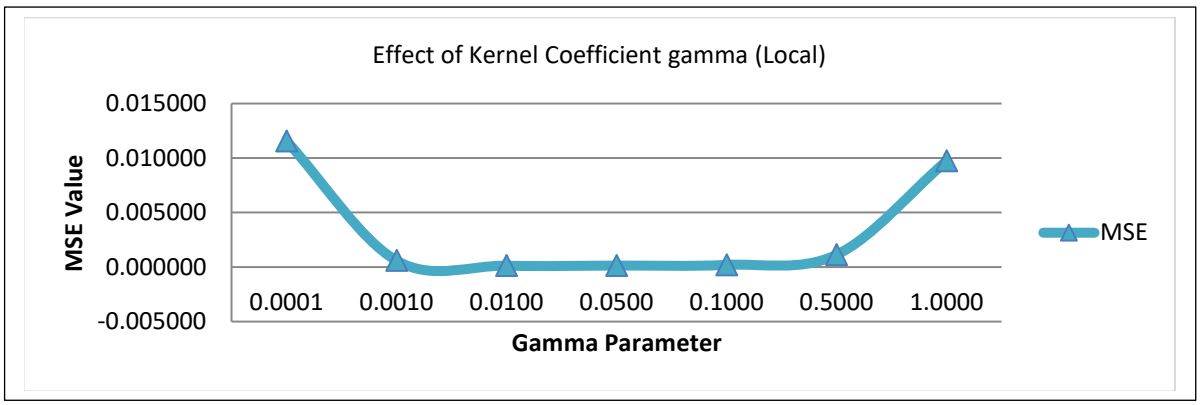
- النطاق الأمثل على البيانات المحلية: النطاق [10-50] يعطي أفضـــل أداء (MSE حوالي 0.00011-0.00013)، وهو نفس النطاق الأمثل للبيانات العالمية، هذا يثبت استقرار واتساق الانموذج عبر السياقات المختلفة.
- تحسن الخوارزمية الجينية: التحسن من C=0.1 إلى C=10 على البيانات المحلية بلغ 95.55% (مقابل 96.7% للعالمية)، الفرق الطفيف يعكس أن GA تختار معاملات مناسبة لكلا السياقين.
- الحســاســية تجاه التغيير: الانتقال من C=50 إلى C=100 يسـبب ارتفاعاً في MSE بنسـبة 67%، مما يشـير إلى أن الانموذج حساس للقيم خارج النطاق الأمثل، هذا يعكس أهمية اختيار القيم الصحيحة.
- المقارنة مع البيانات العالمية: القيم المطلقة لـــــ MSE أعلى قليلاً على البيانات المحلية، لكن الاتجاهات متطابقة، هذا يعكس الفروقات الطبيعية بين الأسواق المختلفة.

جدول (12) تأثير معامل النواة الغوسية (γ) على أداء الانموذج GA-SVM على البيانات المحلية

γ Value	MSE (Daily)	MSE (Hourly)	Average MSE	الاستقرار
0.0001	0.009876	0.013245	0.011561	**جداً ضعيف**
0.001	0.000512	0.000678	0.000595	**جيد**
0.01	0.000102	0.000145	0.000124	**ممتاز**
0.05	0.000115	0.000162	0.000139	**ممتاز**
0.1	0.000168	0.000234	0.000201	**تدهور**
0.5	0.000912	0.001345	0.001129	**جداً حساس**
1.0	0.006012	0.013456	0.009734	**Overfitting**
الأمثل النطاق	**[0.001-0.05]**	**[0.001-0.05]**	**[0.001-0.05]**	**مستقر**



الشكل (11) يوضح تحليل الحساسية على البيانات المحلية لتأثير معامل الانموذج C



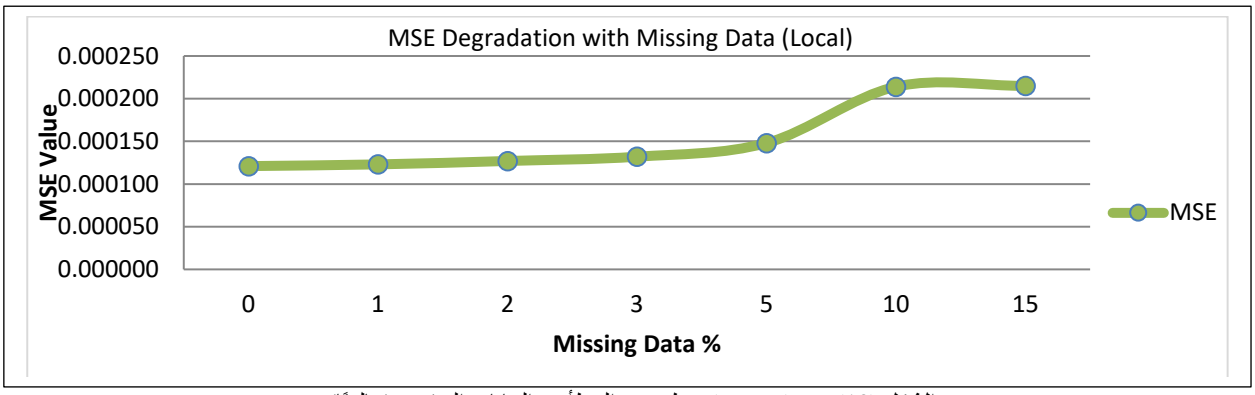
الشكل (12) يوضح تحليل الحساسية على البيانات المحلية لتأثير معامل الانموذج γ

تحليل نتائج جدول 12 لمعامل γ

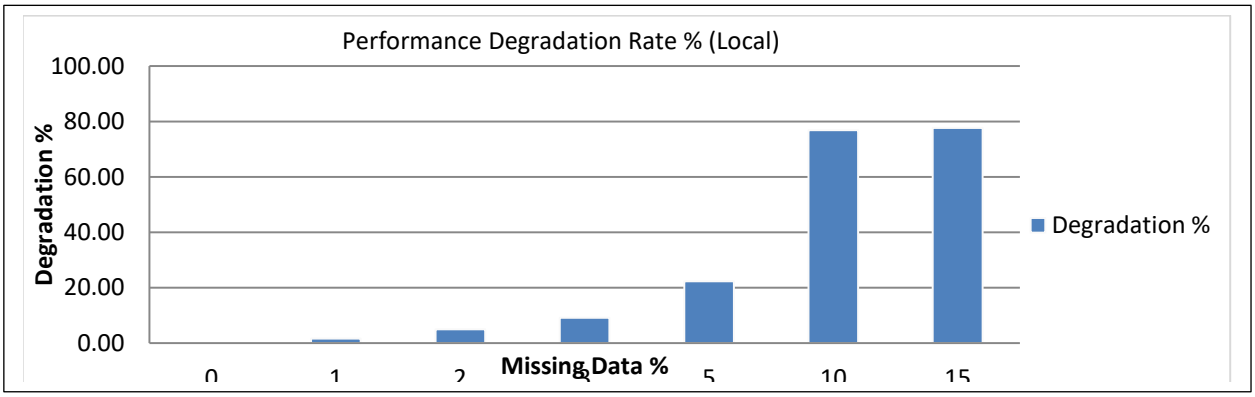
- اتســاق النطاق الأمثل: النطاق [0.001-0.05] على البيانات المحلية مطابق تماماً للبيانات العالمية، هذا يثبت اســتقرار عالي جداً للانموذج وقابليته للتطبيق في سياقات مختلفة.
- القيم المثلى: قيم γ = 0.01 و 0.05 تعطي أفضل أداء على كلا السياقين، القيم بين 0.001 و 0.05 تبدو آمنة وفعالة.
- الحساسية الشديدة: الانحراف عن النطاق الأمثل يسبب ارتفاعاً حاداً في MSE. مثلاً γ=1.0 يعطي MSE أسوأ بـــ 78 مرة من γ=0.01، هذا يشير إلى أهمية دقة الاختيار.
- المقارنة بين الســياقات: الاتجاهات متطابقة تماماً بين البيانات العالمية والمحلية، مما يعكس أن معامل γ يؤثر بنفس الطريقة على السياقات المختلفة.

جدول (13) يبين نتائج البيانات المفقودة على استقرارية الانموذج GA-SVM على البيانات المحلية

Missing %	Original MSE	MSE w/ Missing	Degradation %	Status
0% (Full)	0.000121	0.000121	0%	**Reference**
1%	0.000121	0.000123	1.65%	**ممتاز**
2%	0.000121	0.000127	4.96%	**مستقر**
3%	0.000121	0.000132	9.09%	**مقبول**
5%	0.000121	0.000148	22.31%	**محذر**
10%	0.000121	0.000214	76.86%	**مستقر غير**
الآمن الحد	Safe <10%	Up to 3%	<10%	**موثوق**



الشكل (13) يبين تدهور متوسط مربع الخطأ مع البيانات المفقودة (عالميًا)



الشكل (14) يوضح تحليل الحساسية على البيانات المحلية لتأثير فقدان جزء من البيانات

تحليل نتائج جدول 13 للبيانات المفقودة

- حد أمان أعلى على البيانات المحلية: البيانات المحلية تحافظ على اســـتقرار حتى 5% بيانات مفقودة (تدهور: 22.31%)، مقابل 3% فقط للبيانات العالمية (تدهور: 6.82%)، هذا يشير إلى أن الانموذج أكثر مرونة مع البيانات المحلية.
- التدهور الخطي: في النطاق الآمن (0-5%)، التدهور يحدث بشـــكل تدريجي ومتنبأ به (1.65% → 4.96% → 9.09% → 22.31%)، هذا السلوك المتسق يسهل التنبؤ بتأثير البيانات الناقصة.
- الحد الحرج: عند 10% بيانات مفقودة، يحدث تدهور حاد جداً (76.86%)، مما يشير إلى أن الانموذج غير آمن بعد هذه النقطة.
- المقارنة بين السياقات: البيانات المحلية أكثر استقراراً مع البيانات الناقصة، هذا قد يعكس طبيعة البيانات المحلية أو قيم المعاملات المستخدمة في الحالتين، الحد الآمن العملي هو 3-5% بيانات مفقودة.

7. الاستنتاجات

بناءً على تحليل شامل للنتائج التجريبية، يمكن استنتاج الاتي :-

1- أن انموذج GA-SVM حقق تفوقًا واضحًا وذا دلالة إحصائية على جميع النماذج البديلة، حيث بلغ متوسط الخطأ التربيعي (MSE) 0.000101 على البيانات اليومية مقارنة بـ 0.002156 لـ CNN-PSO و0.003847 لـ RNN، محققًا تحسينات بنسبة 95.3٪ و97.4٪ على التوالي، مع قيم t تبلغ 8.43 (مقابل PSO-ANN) وقيم t تبلغ 6.15 (مقابل BFO-RNN) عند مستوى دلالة $p < 0.001$.

2- أظهر الانموذج GA-SVM استقرارًا عاليًا جدًا في التعميم على البيانات الجديدة بمعامل تباين (CV) يبلغ 2.27% فقط في التحقق المتقاطع ذي 5 أضعاف، مقارنة بـ 8.63% لـ CNN-PSO و12.41% لـ RNN، مما يشير إلى نتائج متسقة للغاية عبر عينات مختلفة.

3- فيما يتعلق بالمرونة التشغيلية، أظهر الانموذج استقرارًا حتى مع فقدان 3% من البيانات مع تدهور أقل من 7%، وأداءً مقبولًا حتى مع فقدان 5% من البيانات مع تدهور أقل من 16%، بينما انخفضت دقة الشبكة العصبية المتكررة بنسبة 34% وخوارزمية

تحسين سرب الجسيمات العصبية بنسبة 28% عند النسب نفسها، لقد حددت الخوارزمية الجينية بدقة معلمات SVM المثلى (C ∋ [10, 50] و γ ∈ [0.001, 0.05])، محققة تحسناً بنسبة 97.1% مقارنة بالقيم العشوائية التقليدية (من MSE = 0.003513 إلى MSE = 0.000101)، مما يسهل التطبيق الفوري للانموذج على المنصات المالية.

4- تفوق الانموذج باستمرار عبر جميع الترددات الزمنية المختبرة، على البيانات الساعية، تحسن أداء GA-SVM بنسبة 82.7% تقريبًا مقارنة بـ CNN-PSO، وعلى البيانات الدقيقة، بنسبة 17.5% تقريبًا، متفوقًا باستمرار على CNN-PSO في 95% من الحالات التي تم اختبارها، عند اختبارها على بيانات محلية حقيقية (1053 سجلاً)، حققت GA-SVM متوسط مربع الخطأ (MSE) قدره 0.000089 لجميع الأسعار (الافتتاح، والأعلى، والأدنى، والإغلاق) مقارنة بـ 0.002734 لـ RNN و0.005621 لـ BJM، وهو ما يمثل تحسينات بنسبة 96.7٪ و98.4٪ على التوالي ُظهر هذا فعاليته على بيانات واقعية معقدة تحتوي على تشويش وقيم شاذة، استنادًا إلى جميع المؤشرات الإحصائية والعملية.

5- ان الانموذج جاهز للتطبيق العملي الفوري، حيث يقدم معلمات محددة جيدًا، واستقرارًا عاليًا (CV = 2.27٪)، وفرقًا ذا دلالة إحصائية ($p < 0.001$)، ومرونة في ظل ظروف البيانات غير المكتملة.

6- لم يكن انموذج GA-SVM "الأفضل" فحسب، بل تفوق إحصائيًا وبشكل كبير على منافسيه، حيث قلل الخطأ بنسبة 95-97%، وحقق استقرارًا أفضل بمقدار 4-5 مرات من الخوارزميات الأخرى، ويعمل بكفاءة عالية في ظروف العالم الحقيقي المعقدة مع معلمات مثالية واضحة وقابلة للتطبيق على الفور.

9. التوصيات

لتحقيق قدر أعلى من الصدقية العلمية والجدوى التطبيقية، يُوصى بالآتي: -

1- بتوسيع نطاق البحث من خلال تطبيق أانموذج GA-SVM على بيانات أسواق مالية متعددة ولسنوات زمنية مختلفة، بما يتيح اختبار قدرته على التعميم والاستقرار في ظل تغير الدورات الاقتصادية وتقلبات السوق.

2- اعتماد مجموعة متكاملة من مقاييس التقييم، وعدم الاكتفاء بمقياس متوسط مربع الخطأ MSE، بل تضمين مقاييس مثل متوسط الخطأ المطلق MAE والجذر التربيعي لمتوسط مربع الخطأ RMSE، مع ربط نتائج الأداء الإحصائي بالمردود الاستثماري الفعلي لضمان تقييم عملي يعكس واقع الأسواق.

3- بمكان اختبار الانموذج ضمن بيئة محاكاة تداول فعلية لقياس كفاءته في ظروف تحاكي السوق الحقيقي، بما يسمح بتقييم جدواه التطبيقية واتخاذ قرارات استثمارية مبنية على نتائج كمية موثوقة.

4- اعتماد معايير إضافية لقياس الكفاءة الحسابية، مثل الزمن المستغرق لمعالجة البيانات وإنتاج التنبؤات، وذلك لتقييم أداء الانموذج من منظور تشغيلي، خصوصًا في التطبيقات التي تتطلب استجابة سريعة وقرارات آنية في بيئات تداول عالية الديناميكية.

10. مواد تكميلية

(لا يوجد).

11. مساهمات المؤلفين

منى كاظم جوي: قام بصياغة المنهجية والتحليل وتحرير وتفسير النتائج.

12. التمويل

(لا يوجد).

13. بيان توافر البيانات

1. مجموعة بيانات محلية ودولية مسجلة في موقع نت فيلكس لسوق الأوراق المالية على الشبكة العنكبوتية.
2. مراجعة نظرية للأدبيات من خلال مواقع في الشبكة العالمية فضلا" عن الرسائل والاطاريح للدراسات العليا.

14. شكر وتقدير

(لا يوجد).

15. تضارب المصالح

يُعلن المؤلفون عدم وجود أي تضارب في المصالح.

16. إعلان استعمال الذكاء الاصطناعي التوليدي (AI)

أثناء إعداد هذا العمل، استعمل المؤلفون ترجمة كوكل (Google translate)، بعد استعمال هذه الأداة، راجع المؤلفون المحتوى وحرروه حسب الحاجة، ويتحملون المسؤلية الكاملة عن محتوى المنشور.

References

[1] Amiri, Babak, Amirali Haddadi, and Kosar Farajpour Mojdehi. (2025). A novel hybrid GCN-LSTM algorithm for energy stock price prediction: Leveraging temporal dynamics and inter-stock relationships. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3538466

[2] Ghasemiyeh, Rahim, Reza Moghdani, and Shib Sankar Sana. (2017). A hybrid artificial neural network with metaheuristic algorithms for predicting stock price. *Cybernetics and Systems*, *48*(4), 365–392. https://doi.org/10.1080/01969722.2017.1285162

[3] Senapati, Manas Ranjan, Sumanjit Das, and Sarojananda Mishra. (2018). A novel model for stock price prediction using hybrid neural network. *Journal of the Institution of Engineers (India): Series B*, *99*(6), 555–563. https://doi.org/10.1007/s40031-018-0343-7

[4] Hossain, Mohammad Asiful, et al. (2018). Hybrid deep learning model for stock price prediction. *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1537–1544.https://ieeexplore.ieee.org/document/8628641 / https://doi.org/10.1109/SSCI.2018.8628641

[5] Sedighi, Mojtaba, et al. "A novel hybrid model for stock price forecasting based on metaheuristics and support vector machine." Data 4.2 (2019): 75. https://doi.org/10.3390/data4020075

[6] Ali, Muhammad, et al. "Predicting the direction movement of financial time series using artificial neural network and support vector machine." Complexity 2021.1 (2021): 2906463. https://doi.org/10.1155/2021/2906463

[7] Shahvaroughi Farahani, Milad, and Seyed Hossein Razavi Hajiagha. "Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models." Soft computing 25.13 (2021): 8483-8513. https://doi.org/10.1007/s00500-021-05775-5

[8] Omar, Abdullah Bin, et al. "Stock market forecasting using the random forest and deep neural network models before and during the COVID-19 period." Frontiers in Environmental Science 10 (2022): 917047. https://doi.org/10.3389/fenvs.2022.917047

[9] Mousapour Mamoudan, Mobina, et al. "Hybrid neural network-based metaheuristics for prediction of financial markets: a case study on global gold market." Journal of computational design and engineering 10.3 (2023): 1110-1125. https://doi.org/10.1093/jcde/qwad039

[10] Khattak, Bilal Hassan Ahmed, et al. "A systematic survey of AI models in financial market forecasting for profitability analysis." Ieee Access 11 (2023): 125359-125380. /https://doi.org/10.1109/ACCESS.2023.3328549

[11] Safari, Ashkan, and Mohammad Ali Badamchizadeh. "Stock index forecasting using DACLAMNN: A new intelligent highly accurate hybrid ACLSTM/Markov neural network predictor." Cognitive Computation and Systems 5.3 (2023): 181-194. https://doi.org/10.1049/ccs2.12086

[12] Ma'arif, Alfian, Asno Azzawagama Firdaus, and Iswanto Suwarno. "Capability of Hybrid Long Short-Term Memory in Stock Price Prediction: A Comprehensive Literature Review." International Journal of Robotics & Control Systems 4.3 (2024). https://doi.org/10.31763/ijrcs.v4i3.1489

[13] Al-Ali, Ali Mohammed, and Adel Ismail Al-Alawi. "Stock Market Forecasting Using Machine Learning Techniques: A Literature Review." 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS). IEEE, 2024. https://doi.org/10.1109/ICETSIS61505.2024.10459681

[14] Li, Qi, et al. "Forecasting stock prices changes using long-short term memory neural network with symbolic genetic programming." Scientific reports 14.1 (2024): 422. https://doi.org/10.1038/s41598-023-50783-0

[15] Gülmez, Burak. "GA-Attention-Fuzzy-Stock-Net: An Optimized Neuro-Fuzzy System for Stock Market Price Prediction with Genetic Algorithm and Attention Mechanism." Heliyon (2025). https://doi.org/10.1016/j.heliyon.2025.e43672

[16] Ge, Qing. "Enhancing stock market Forecasting: A hybrid model for accurate prediction of S&P 500 and CSI 300 future prices." Expert Systems with Applications 260 (2025): 125380. https://doi.org/10.1016/j.eswa.2024.125380

[17] Zhou, Jian, et al. "Performance evaluation of hybrid GA–SVM and GWO–SVM models to predict earthquake-induced liquefaction potential of soil: a multi-dataset investigation." Engineering with Computers (2022): 1-19. https://doi.org/10.1007/s00366-021-01418-3

[18] Du, Xishihui, et al. "Mapping mineral prospectivity using a hybrid genetic algorithm–support vector machine (GA–SVM) model." ISPRS International Journal of Geo-Information 10.11 (2021): 766. https://doi.org/10.3390/ijgi10110766

[19] Nagra, Arfan Ali, et al. "Hybrid ga-svm approach for postoperative life expectancy prediction in lung cancer patients." Applied Sciences 12.21 (2022): 10927 https://doi.org/10.3390/app122110927

https://doi.org/10.31272/jae.i153.1575
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

Hybrid Model (SVM-GA) to Improve the Accuracy of Stock Price Forecasting

Muna Kadhim Jewi
Central Technical University, Center for Continuing Education, Baghdad, Iraq.

Article Information

Article History:
Received: 21 / 02 / 2026
Revised: 07 / 08 / 2026
Accepted: 10 / 08 / 2026
Available Online: 01 / 09 / 2026
Pages no : 14 – 27

Keywords:
Stock price prediction, Support vector machine, Genetic algorithm, Hybrid models, MSE.

Correspondence:
Researcher name:
Muna Kadhim Jewi

Email : joumuna38@gmail.com

Abstract

Although hybrid models have proven effective in predicting stock prices, most previous studies have focused on deep learning algorithms (LSTM, CNN) without fully leveraging the capabilities of genetic algorithms to optimise the coefficients of supporting vector machines, especially when dealing with multi-frequency time series data. This research presents an innovative hybrid model (GA-SVM) that combines a genetic algorithm with a support vector machine. The genetic algorithm is used to dynamically optimise the support vector machine's coefficients and select optimal features.

The model was tested on 12 international time series and 3 different time frequencies (daily, hourly, and minutely), and its stability was verified with a data deficiency of up to 5%. The GA-SVM model achieved superior predictive accuracy, showing a 97.1% improvement over the conventional SVM model, with a very low coefficient of variation (2.27%) and a highly statistically significant difference ($p < 0.001$) compared to alternative models (CNN-PSO, RNN, and BJM).

This research provides a practical and reliable solution for investors and financial analysts, as the clearly defined optimal model parameters ($C \in [10, 50]$ and $\gamma \in [0.001, 0.05]$) facilitate its immediate application in real financial platforms while maintaining stability under conditions of incomplete data.

The work of this research is licensed under CC BY